Computer Science Department

TECHNICAL REPORT

"A Lower Bound on the Distance to the Nearest Uncontrollable System"

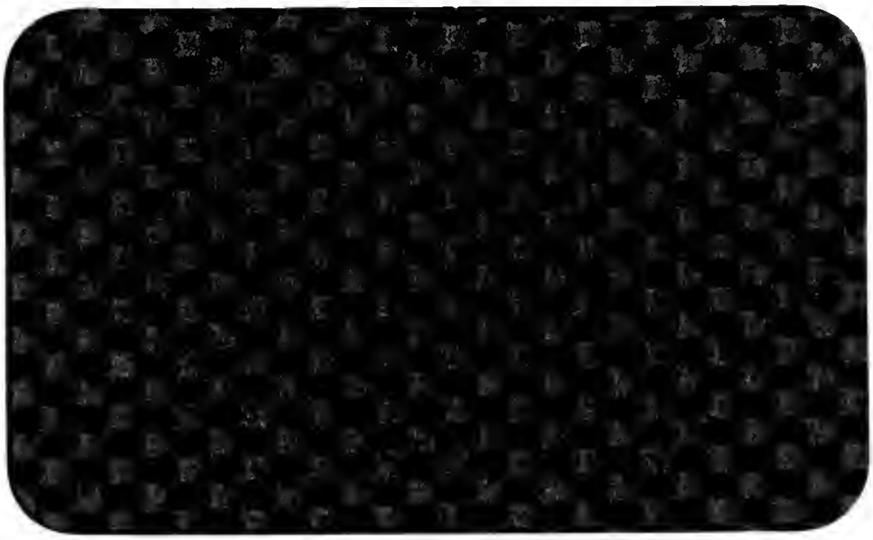
J. Demmel

Technical Report #272

January 1987

NEW YORK UNIVERSITY

Department of Computer Science
Courant Institute of Mathematical Sciences
251 MERCER STREET, NEW YORK, N.Y. 10012



"A Lower Bound on the Distance to the Nearest Uncontrollable System"

J. Demmel

Technical Report #272

January 1987

A Lower Bound on the Distance to the Nearest Uncontrollable System

James W. Demmel
Computer Science Department
Courant Institute of Mathematical Sciences
251 Mercer Str.
New York, NY 10012
(212)460-7494

Abstract

One measure of the controllability of a linear system is the distance from the pair (A,B) to the nearest uncontrollable pair. We provide an easily computable and nearly attainable lower bound on this distance in terms of the "staircase form" of the pair (A,B).

Keywords: Controllability

Introduction

Let A be an n by n matrix and B an n by m matrix. The pair (A,B) is called *controllable* if the matrix $C(A,B) \equiv [B \,|AB\, |A^2B| \cdots |A^{n-1}B]$ has full rank; otherwise it is uncontrollable. One measure of the controllability of (A,B) is the *distance* from (A,B) to the nearest uncontrollable pair $(A+\delta A, B+\delta B)$: the distance is

$$\|\delta A,\delta B\|_F \equiv (\sum_{ij} |\delta A_{ij}|^2 + \sum_{ij} |\delta B_{ij}|^2)^{1/2}$$

if $(A+\delta A, B+\delta B)$ is uncontrollable and $(A+\delta A', B+\delta B')$ is controllable whenever $\|\delta A',\delta B'\|_F < \|\delta A,\delta B\|_F$. We will denote this distance by dist(A,B). dist(A,B) indicates how much A and B may be perturbed without making the system uncontrollable.

Several other workers have attempted to characterize dist(A,B). Eising [3] has shown that

$$\text{dist}(A,B) \equiv \inf_{\lambda} \sigma_{\min}[B\, |A-\lambda I] \tag{1}$$

where $\sigma_{\min}(X)$ is the smallest singular value of X. This result shows immediately that dist(QAQ^*,QB)=dist(A,B) for any unitary matrix Q. Boley and Lu [2] have investigated the "staircase form" of the system (A,B), i.e. a transformation to an equivalent system (QAQ^*,QB) where Q is a unitary matrix, and in case B has only a single column, QB has only its first entry nonzero and $QAQ^*=H$ is upper Hessenberg. If B has more than one column there is a "block" version of this form: at most the first m rows of QB are nonzero and QAQ^* is a conforming block upper Hessenberg matrix (more on this below). Van Dooren [5] describes an algorithm for computing this staircase form. This algorithm uses the norms of the subdiagonal blocks $\|H_{i,i-1}\|$ to determine whether (A,B) is controllable. Note that by setting any $H_{i,i-1}$ to zero (H,QB) becomes uncontrollable, so that $\min_i \|H_{i,i-1}\|$ is an upper bound on dist(A,B). Boley and Lu [2] have shown that $\min_i \|H_{i,i-1}\|$ can severely overestimate dist(A,B), and have used this fact to criticize the staircase algorithm as a means of estimating dist(A,B).

In this note we give a lower bound on dist(A,B) in terms of the product of the $\sigma_{\min}(H_{i,i-1})$. We show by means of an example that this bound is nearly attainable. Thus the staircase algorithm can be used to give a lower bound on dist(A,B) as well as an upper bound.

Main Result

We prove the theorem in detail when B consists of a single column; the multiinput case is analogous.

Theorem 1: Suppose B is a single column. Assume without loss of generality that $||B||=1$, $||A|| \equiv \sup_{x\neq 0} ||Ax||/||x|| = 1$, and that (A,B) is in staircase form as defined above. Let $p \equiv \prod_{i=2}^{n} |A_{i,i-1}|$ be the magnitude of the product of the subdiagonal entries of A. Then

$$\text{dist}(A,B) \geq \frac{p}{n\cdot(b(n))^n}$$

where

$$b(n) \equiv \frac{1}{2}\cdot\left(\left(3+\frac{1}{n}\left(\frac{2}{3+5^{1/2}}\right)\right)^n + \left[\left(\left(1+\frac{1}{n}\left(\frac{2}{3+5^{1/2}}\right)\right)^n\right)^2 + 4\right]^{1/2}\right)$$

Remark: $b(2)\approx 2.672$ and quickly decreases towards its limit $(3+5^{1/2})/2 \approx 2.618$ as n increases.

Proof: By choosing a diagonal matrix D with 1s and -1s on the diagonal, DAD can have subdiagonal entries with arbitrary sign, so we assume without loss of generality that all $A_{i,i-1}>0$. Similarly, we can assume that $B_1=1$. Eising's result (1) implies that dist$(A,B) = \sigma_{\min}[B\,|A-\lambda I]$ for some fixed λ. To bound this quantity from below, we need the following fact: $\sigma_{\min}[X\,|Y] \geq \sigma_{\min}[X]$. This follows from the fact that the eigenvalues of $[X\,|Y]\cdot[X\,|Y]^* = XX^*+YY^*$ are all at least as large as the eigenvalues of XX^* [4]. Now let $[X\,|Y] = [B\,|A-\lambda I]$ where X consists of B and the first $n-1$ columns of $A-\lambda I$. Thus dist$(A,B) \geq \sigma_{\min}(X) = ||X^{-1}||^{-1}$, and we need to bound $||X^{-1}||$ from above, which we do as follows.

X is upper triangular with $X_{11}=1$ and $X_{jj}=A_{j,j-1}$. The superdiagonal entries are bounded by $|X_{j,j+1}| = |A_{jj}-\lambda| \leq 1+|\lambda|$ and $|X_{i,j}| = |A_{i,j-1}| \leq 1$ for $j>i+1$. It is tedious but straightforward to show that if we construct the matrix Z where

$$Z_{ij} \equiv \begin{cases} 0 & \text{if } i>j \\ X_{ij} & \text{if } i=j \\ -1-|\lambda| & \text{if } i=j-1 \\ -1 & \text{if } i<j-1 \end{cases}$$

then Z^{-1} is a nonnegative upper triangular matrix with $Z_{ij}^{-1} \geq |X_{ij}^{-1}|$. This implies $||Z^{-1}|| \geq ||X^{-1}||$. Since $||Z^{-1}|| \leq n\cdot\max_{ij} |Z_{ij}^{-1}|$, it suffices to bound the largest entry in Z^{-1}. Again, it is tedious but straightforward to show that the largest entry in Z^{-1} is Z_{1n}^{-1}, the entry in the upper right corner.

For convenience let $s_i \equiv A_{i+1,i}$, $s_0=1$, and $a = 1+|\lambda|$. We can define the following recurrence for the entries of the last column of Z^{-1}:

$$Z_{nn}^{-1} = s_{n-1}^{-1}$$
$$Z_{n-1,n}^{-1} = s_{n-2}^{-1}(a \cdot Z_{nn}^{-1})$$
$$Z_{jn}^{-1} = s_{j-1}^{-1}(a \cdot Z_{j+1,n} + Z_{j+2,n} + \cdots + Z_{nn})$$

Letting $Y_j \equiv Z_{jn}^{-1} \cdot (s_{n-1} \cdots s_{j-1})$ we may rewrite this recurrence as

$$Y_n = 1$$
$$Y_{n-1} = a$$
$$Y_j = aY_{j+1} + s_j Y_{j+2} + s_j s_{j+1} Y_{j+3} + \cdots + s_j \cdots s_{n-2} Y_n$$

Since $0 < s_i \le 1$, Y_j is bounded above by $\bar{Y}_j$ where

$$\bar{Y}_n = 1$$
$$\bar{Y}_{n-1} = a$$
$$\bar{Y}_j = a\bar{Y}_{j+1} + \bar{Y}_{j+2} + \bar{Y}_{j+3} + \cdots + \bar{Y}_n$$

We can solve this last recurrence explicitly

$$\bar{Y}_j = \left(\frac{a-1+(a^2-2a+5)^{1/2}}{2 \cdot (a^2-2a+5)^{1/2}} \right) \cdot \left(\frac{a+1+(a^2-2a+5)^{1/2}}{2} \right)^{n-j} + \left(\frac{1-a+(a^2-2a+5)^{1/2}}{2 \cdot (a^2-2a+5)^{1/2}} \right) \cdot \left(\frac{a+1-(a^2-2a+5)^{1/2}}{2} \right)^{n-j}$$

Since $a = 1 + |\lambda| \ge 1$, the first term in this formula dominates the second so

$$Z_{1n}^{-1} = \frac{Y_1}{p} \le \frac{\bar{Y}_1}{p} \le \frac{2}{p} \cdot \left(\frac{a-1+(a^2-2a+5)^{1/2}}{2 \cdot (a^2-2a+5)^{1/2}} \right) \cdot \left(\frac{a+1+(a^2-2a+5)^{1/2}}{2} \right)^{n-1}$$
$$\le \frac{1}{p} \cdot \left(\frac{a+1+(a^2-2a+5)^{1/2}}{2} \right)^n$$

or

$$\text{dist}(A,B) \ge \frac{1}{nZ_{1n}^{-1}} \ge \frac{p}{n} \cdot \left(\frac{2}{|\lambda|+2+(|\lambda|^2+4)^{1/2}} \right)^n \equiv b_1(|\lambda|) \qquad (2)$$

This bound decreases as $|\lambda|$ increases, so we need another bound that increases as $|\lambda|$ increases in order to get a bound independent of $|\lambda|$. To do this we use the same trick as before: $\text{dist}(A,B) = \sigma_{\min}[B\,|A-\lambda I] \ge \sigma_{\min}[A-\lambda I]$. As before, we need to bound $||(A-\lambda I)^{-1}||$ from above to bound $\text{dist}(A,B)$ from below. If $|\lambda| > 1$, then since $||A|| = 1$

$$||(A-\lambda I)^{-1}|| = ||\lambda^{-1}(I - A/\lambda)^{-1}|| \le |\lambda^{-1}| \;/\; (1 - ||A/\lambda||) = 1 \;/\; (|\lambda| - 1)$$

Thus we have our second lower bound on $\text{dist}(A,B)$

$$\text{dist}(A,B) \ge |\lambda| - 1 \equiv b_2(|\lambda|) \quad \text{if} \quad 1 \le |\lambda| \ .$$

The lower bound $b_1(|\lambda|)$ is a decreasing function of $|\lambda|$ and $b_2(|\lambda|)$ is an increasing function of $|\lambda|$. Since $b_2(1) < b_1(1)$ and $b_2(2) > b_1(2)$, $b_1(\lambda_1) = b_2(\lambda_1)$ for some $1 \le \lambda_1 \le 2$.

Solving for λ_1 we get

$$b_1(\lambda_1) = \frac{p}{n} \left(\frac{2}{\lambda_1 + 2 + (\lambda_1^2+4)^{1/2}} \right)^n = \lambda_1 - 1 = b_2(\lambda_1) \ .$$

Since $b_2(\lambda_1)$ has slope 1, λ_1 is clearly bounded above by $1+b_1(1) \le 1+n^{-1}(2/(3+\sqrt{5}))^n \equiv \lambda_2$. This implies that

$$\mathrm{dist}(A,B) \ge b_1(\lambda_2) = \frac{p}{n}\cdot\left(\frac{2}{3+\frac{1}{n}\left(\frac{2}{3+5^{1/2}}\right)^n + \left[\left(1+\frac{1}{n}\left(\frac{2}{3+5^{1/2}}\right)^n\right)^2+4\right]^{1/2}}\right)^n$$

as desired. □

Now we give an example that shows that $\mathrm{dist}(A,B)$ can in fact be as small as p. Let B be the first column of the n by n identity matrix as before and A bidiagonal with $A_{jj}=-1$, $j<n$, $A_{nn}=0$, and $A_{j,j-1}=\epsilon_j \ll 1$. Then $\|A\| \approx 1$, and $\mathrm{dist}(A,B) \le \sigma_{\min}[B\,|A] = \sigma_{\min}(X)$, where X consists of B and the first $n-1$ columns of A, since the last column of A is zero. X^{-1} is easily calculated and its largest entry X_{1n}^{-1} is seen to be $\prod_j \epsilon_j^{-1} = p^{-1}$, so $\mathrm{dist}(A,B) \le p$ as desired. Also, Boley [1] gives an example where $\mathrm{dist}(A,B)$ decreases as 2^{-n} with constant p, so the factor $b(n)^{-n}$ in the bound, while not necessarily attainable, does reflect the behavior of the problem. If either the minimizing λ in (1) is close to zero, or if each subdiagonal entry $A_{i+1,i}$ is very small, one can show that $b(n)^{-n}$ can be replaced by a factor close to 2^{-n}. Note also that when p is small λ_1 in the proof is close to 1 and the $b(n)^{-n}$ factor approaches its limit 2.618^{-n}.

To state the general result, we need to define staircase form. Suppose

$$B = \begin{bmatrix} B_1 \\ 0 \\ \cdot \\ 0 \end{bmatrix} \quad \text{and} \quad A = \begin{bmatrix} A_{11} & \cdot & \cdot & A_{1,j} \\ A_{21} & \cdot & \cdot & \cdot \\ 0 & \cdot & \cdot & \cdot \\ \cdot & 0 & A_{j,j-1} & A_{j,j} \end{bmatrix} \qquad (3)$$

where A is block upper Hessenberg as shown, B_1 has as many rows as A_{11}, B_1 has full row rank, and each subdiagonal block $A_{i,i-1}$ has full row rank. Then we say (A,B) is in staircase form. As mentioned above, there is a unitary matrix Q that transforms any pair (A,B) into staircase form (QAQ^*,QB) [5].

Theorem 2: Let B be an n by m matrix and A an n by n matrix. Assume without loss of generality that $\|A\|=\|B\|=1$ and that (A,B) is in staircase form with j blocks as in (3). Let $p=\sigma_{\min}(B_1)\cdot\prod_{i=2}^{j}\sigma_{\min}(A_{i,i-1})$. Let $b(\cdot)$ be defined as in the statement of Theorem 1. Then

$$\mathrm{dist}(A,B) \ge \frac{p}{j\cdot(b(j))^j} \ .$$

Since generically $j=\lceil n/m \rceil$, this lower bound is an increasing function of m (for fixed p).

Conclusions

We have shown that the staircase form of a single input control system can be used to provide a lower as well as an upper bound on the distance to the nearest uncontrollable system. This bound is easily computed from the staircase form, and is shown to be nearly attainable for some systems. The gap between the upper and lower bounds can be very large, and so the problem of accurately and inexpensively estimating the distance to the nearest uncontrollable system remains open.

References

[1] D. L. Boley, "Computing rank-deficiency of rectangular matrix pencils," TR 86-14, Computer Science Dept., University of Minnesota, Minneapolis, April 1986

[2] D. L. Boley, W. S. Lu, "Measuring How far a controllable system is from an uncontrollable one," IEEE Trans. Aut. Contr., Vol. AC-31, No. 3, pp 249-252, March 1986

[3] R. Eising, "Between controllable and uncontrollable," Systems and Control Letters, Vol. 4, No. 5, pp 263-264, July 1984

[4] B. Parlett, "The Symmetric Eigenvalue Problem," Englewood Cliffs, Prentice-Hall, 1980, chap. 10

[5] P. Van Dooren, "The Generalized Eigenstructure Problem in Linear System Theory," IEEE Trans. Aut. Contr., Vol. AC-26, No. 1, pp 111-128, February 1981

NYU COMPSCI TR-272 c.2
Demmel, James
A lower bound on the distance to the nearest uncontrollable system